The Effect of Document Sources on the Evaluation of Embedded Sources and Their Claims When Reading Health Information

Franco Londra, Gastón Saux and Magalí Martínez

July 27, 2022

The Effect of Document Sources on the Evaluation of Embedded Sources and Their Claims When Reading Health Information

Introduction

Although the quality of health information on the internet may be biased or inaccurate (e.g. Dinh et al., 2022; Ng et al., 2021), most adults use the internet to find health information (Lim et al., 2022). In this context, while internet may provide valid and up to date information, being able to critically think about what is read on the web is essential to overcome misinformation (Braasch & Graesser, 2020).

The Content-Source Integration model (CSI; Stadtler & Bromme, 2014) proposes that, if a conflict between texts is detected, motivated readers will try to determine which information to adopt by asking what is true (i.e. first-hand evaluations) or whom to believe (i.e. second-hand evaluations). Laypeople could find first-hand evaluations taxing, as specific domain knowledge may be lacking. Second-hand evaluation could provide a bypass in which evaluations are made via source parameters. Paying attention to sources could facilitate adopting accurate information and neglecting inaccurate information.

Strømsø et al. (2013) pointed out that the representation of multiple sources may be organized in layers. Document Sources (DS) could be thought of as a first layer of information. They have been defined as metadata about a particular document (e.g. about its publisher, date, etc.). Embedded Sources (ES) have been defined as sources cited within a document as the origin of an argument or statement (Bråten et al., 2016). The organization of sources into layers may have an impact on the way readers evaluate them, although evidence about this specific issue is scarce. While some studies have found that people pay attention and use information from DS and ES (Strømsø et al., 2013; Strømsø & Bråten, 2014) others have reported that attention to and use of both is very low (Bråten et al., 2016; Salmerón et al., 2018; Salmerón et al., 2020). Earlier works found that students may have difficulties differentiating layers of source information,

evaluating DS and ES on equal terms, disregarding the fact that one layer is embedded in the other (Strømsø et al., 2013; Strømsø & Bråten, 2014). Nonetheless, more research is needed regarding the potential relationship between source layers and how they are represented.

The Present Study

The objective of this work was to determine if the perceived trustworthiness in DS influences the trustworthiness of ES (H1) and the evaluation of their content statements (H2) when reading about unfamiliar health issues. Additionally, reading times of ES and their statements were also measured to explore a potential effect of DS.

Methodological overview

Participants

104 psychology undergraduates from a South American university (age M = 25.01; SD = 8.26) participated voluntarily. .

Design and materials

DS trustworthiness (trustworthy vs untrustworthy) was manipulated as a single within-subject factor. Two texts were constructed about treatments for a rare genetic disease and DS information was added in a heading. Texts were similar in length (175 vs 160 words) and readability ("Somewhat difficult", Fernández, 1959). Texts were segmented into sections. Section 1 presented the DS. Sections 2 and 6 functioned as introduction and closure. Section 3 introduced an ES which recommended a treatment. Section 4 and 5 presented and described the treatment, respectively. ES were counterbalanced so that they could be read in the context of trustworthy or untrustworthy DS. Table 1 presents an example of one of the final texts. Additionally, texts were presented on screen section by section, with the moving window technique to measure reading times.

Table 1

Example of one of the final texts and its sections

Text example	Sections and content
Magical Nature Magazine **Bi-monthly publication of the** **"Sisters of Gaia Argentina Society"**	Section 1: Document source information (Untrustworthy condition)
Cystic fibrosis is a hereditary disease that mainly affects the lungs, and to a lesser extent the pancreas, liver, and intestine, causing the accumulation of a thick and sticky mucus in these areas. It is one of the most common types of chronic lung disease in children and young adults and it could be life-threatening as patients are vulnerable to pulmonary infections.	Section 2: Introduction
According to Oscar Prat, a nurse in a pulmonology service,	Section 3: Embedded source information
adapted diets are the best treatment to cope with Cystic Fibrosis.	Section 4: Treatment
Patients with this disease need to eat nutritious foods more regularly throughout the day. Kids should eat a lot more calories and fat. This requires learning about the nutrients in each food. With a good diet, it is not necessary to take pills.	Section 5: Treatment's description
In our country, it is estimated that every 7,200 births 1 suffers from the disease, but numbers would increase considerably if the cases in which the disease is not detected are taken into account, so parents and health professionals are encouraged to learn about Cystic Fibrosis and other rare diseases.	Section 6: Closing

Procedure

The experiment was conducted on synchronous sessions of approximately 30 participants. After signing an informed consent they were asked to fill in socio-demographic information. Next, a 20 mins. pre-training was conducted on the importance of source information when reading conflicting information online (adapted from Perez et al., 2018). Then, participants were instructed to read the two experimental texts and to write an essay stating which treatment was the best, as part of a fictitious scholarly task. Finally, participants were asked to evaluate the trustworthiness of each DS and ES on a 10-points scale.

Statistical Analyses

The effect of DS trustworthiness on ES trustworthiness (H1) was analyzed by fitting a linear mixed model (LMM). The effect of DS trustworthiness on treatment selection (H2) was analyzed with chi-square tests. Last, to assess the potential influence of DS on reading times, generalized linear mixed models (GLMM) were fitted. Participants ID and Texts were added as sources of random variance. Prior self-reported knowledge was added to the models as control.

Key findings

First, trustworthy DS increased mean trustworthiness scores of ES by 1.33, t(205)= 3.32, p=.001, $95\%CI$ [0.55, 2.11]. Mean perceived trustworthiness of ES was 5.91 (SD= 2.75) under trustworthy DS and 4.59 (SD= 3.03) under untrustworthy DS.

Second, the treatment proposed by the ES under trustworthy DS was selected as the best more often than the one proposed by the ES under untrustworthy DS, 71.58% vs 28.42%, $\chi2$ (1) = 31, p<.001.

Last, information about the DS and the proposed treatment were read for longer when presented by an untrustworthy DS. On average people took 2.11 more seconds to read the section presenting the DS ($CI95\%$ [0.69, 3.53], t=-2.94 p= .003) and 1.54 more seconds ($CI95\%$ [0.1, 2.98], t=-2.12, p=.034) to read the section presenting the proposed treatment in the context of an untrustworthy DS.

Discussion

As expected, the perceived trustworthiness of ES increased when presented in the context of a trustworthy DS (H1). This finding suggests a layered representation of document and embedded sources (Strømsø, et al., 2013).

We also found that trustworthy DS lead participants to select the treatment proposed by the ES as the best. These results are in line with prior works on the use of source information when dealing with conflicts between texts (Braasch & Kessler, 2021).

Last, we found that DS trustworthiness influenced reading times of the DS itself and of the treatment proposed within the text, with participants taking longer when the DS was untrustworthy. This could be the result of participants' enhanced epistemic vigilance in the untrustworthy condition. Of note, instructions and pre-training explicitly asked participants to evaluate sources. This may have induced them to read strategically, particularly when confronted with an untrustworthy document source. The extent to which these results generalize to situations in which people read less strategically (i.e., less attentive to sources' trustworthiness) is unclear, but results may differ.

References

Braasch, J. L., & Kessler, E. D. (2021). Working toward a theoretical model for source comprehension in everyday discourse. *Discourse Processes*, *58*(5-6), 449-467. https://doi.org/10.1080/0163853X.2021.1905393

Braasch, J. L. G., & Graesser, A. C. (2019). Avoiding and overcoming misinformation on the Internet. In R. J. Sternberg, & D. F. Halpern (Eds.), *Critical thinking in psychology* (2nd ed., pp. 125–151). Cambridge: Cambridge University Press.

Bråten, I., Strømsø, H. I., & Andreassen, R. (2016). Sourcing in professional education: Do text factors make any difference?. *Reading and Writing*, *29*(8), 1599-1628. https://doi.org/10.1007/s11145-015-9611-y

Dinh, T., Flaxman, T., Shea, K., & Singh, S. S. (2022). What Are Patients Reading? Quality Assessment of Endometriosis Information on the Internet. *Journal of Obstetrics and Gynaecology Canada*, *44*(1), 11-20. https://doi.org/10.1016/j.jogc.2021.08.007

Fernández, H. J. (1959). Medidas sencillas de lecturabilidad. Consigna, 214, 29–32.

Lim, M. S., Molenaar, A., Brennan, L., Reid, M., & McCaffrey, T. (2022). Young Adults'
    Use of Different Social Media Platforms for Health Information: Insights From
    Web-Based Conversations. *Journal of medical Internet research*, *24*(1).
    http://doi.org/10.2196/23656

Ng, J. Y., Dzisiak, D. A., & Saini, J. B. (2021). Cannabis for pain: a cross-sectional survey
    of the patient information quality on the Internet. *Journal of Cannabis Research*,
    3(1), 1-17. http://doi.org/10.1186/s42238-021-00093-x

Salmerón, L., Delgado, P., & Mason, L. (2020). Using eye-movement modelling
    examples to improve critical reading of multiple webpages on a conflicting topic.
    Journal of Computer Assisted Learning, *36*(6), 1038-1051.
    https://doi.org/10.1111/jcal.12458

Salmerón, L., Gil, L., & Bråten, I. (2018). Effects of reading real versus print-out versions
    of multiple documents on students' sourcing and integrated understanding.
    Contemporary Educational Psychology, *52*, 25-35.
    https://doi.org/10.1016/J.CEDPSYCH.2017.12.002

Stadtler, M., & Bromme, R. (2014). The content-source integration model: A taxonomic
    description of how readers comprehend conflicting scientific information. In D. N.
    Rapp & J. L. G. Braasch (Eds.), *Processing inaccurate information: Theoretical
    and applied perspectives from cognitive science and the educational sciences*
    (pp. 379–402). MIT Press.

Strømsø, H. I., & Bråten, I. (2014). Students' sourcing while reading and writing from
    multiple web documents. *Nordic Journal of Digital Literacy*, *9*(2), 92-111.
    https://doi.org/10.1080/07370008.2013.769994

Strømsø, H. I., Bråten, I., Britt, M. A., & Ferguson, L. E. (2013). Spontaneous sourcing

among students reading multiple documents. *Cognition and Instruction*, *31*(2),

176-203. https://doi.org/10.1080/07370008.2013.769994